DOCUMENT RESUME

ED 325 507 TM 015 741

AUTHOR Conley, Carolyn S.; Barnes, Laura L. B.
TITLE The Effect of Method of Evaluation on Student Aspiration of Success, Academic Self-Concept and Test Anxiety.
PUB DATE Oct 90
NOTE 26p.; Paper presented at the Annual Meeting of the Midwestern Educational Research Association (12th, Chicago, IL, October 17-20, 1990).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Achievement; *Academic Aspiration; *College Students; Comparative Testing; Criterion Referenced Tests; Higher Education; Norm Referenced Tests; *Self Concept; Student Attitudes; Student Educational Objectives; *Success; *Test Anxiety; Test Format
IDENTIFIERS *Academic Self Concept; Analogies

ABSTRACT

The effects of three methods of evaluation (norm-referenced, regulated criterion-referenced, and negotiated criterion-referenced) on academic self-concept (ASC), test anxiety, and aspiration of success were studied for 73 college students (71 females and 2 males) from a large midwestern research university. Students were given a limited amount of time to complete a series of difficult word analogies (subjects were given 3 minutes to complete each of 5 trials). Those in the norm-referenced group were told that their scores would be based on a comparison with other students. Subjects in the regulated criterion-referenced groups were told that a score of 10 out of 20 for each trial would be necessary in order to be considered proficient on the analogies test. Subjects in the negotiated criterion-referenced group were allowed to state the score that they thought would indicate proficiency. The time limit was imposed so that students' ASC, test anxiety, and aspiration of success could be studied under induced failure conditions. Data were analyzed using split plot analysis of variance; and it was found that the method of evaluation had no significant effect on subjects' ASC, test anxiety, and aspiration of success. Three tables contain the study data. A 26-item list of references is included. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED325507

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

CAROLYN S. CONLEY

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

The Effect of Method of Evaluation on Student Aspiration of Success, Academic Self-Concept and Test Anxiety

Carolyn S. Conley
Laura L. B. Barnes

Oklahoma State University

Mid-Western Educational Research Association
Annual Meeting, Chicago, Illinois

October 18, 1990

Running Head: Evaluation Methods

TM015741

Abstract

This study examined the effect of three methods of evaluation (norm-referenced, regulated criterion-referenced and negotiated criterion-referenced) on academic self-concept, test anxiety, and aspiration of success of 73 college students. Students in this study were given a limited amount of time to complete a series of difficult word analogies. The time limit was imposed so that students' academic self-concept, test anxiety, and aspiration of success could be studied under induced failure conditions. Data were analyzed using split plot Analysis of Variance and it was found that method of evaluation had no significant effect on academic self-concept, test anxiety, and aspiration of success.

Introduction

Although there is much criticism of grades as a means of evaluation in education, their use is widespread as a measure of learning. The typical student experiences years of being judged in relationship to others by comparisons made between classmates and one's own performance. Since failure to perform apparently has a strong negative effect on the individual's attitudes, and logically, repeated failure should discourage learning, it becomes important to use an evaluation method that produces the least severe effects after failure.

It is the student's perception of his/her performance that is important in evaluation. With norm-referenced evaluation the student's perception is based on the relative judgment of others (Popham, 1971). This type of evaluation has met with criticism in terms of arbitrariness in deciding grades, not encouraging individualized learning, and being poor in predicting success in addition to producing excessive competition, negativeness, and dishonesty (Astin, et al., 1967). Others are now criticizing the norm-referenced approach in terms of the affective outcomes of relative evaluations.

Absolute measures (criterion-referenced, competency

testing, mastery learning) are being suggested to correct many of the defects of relative measures (Block, 1971; Block, 1974; Bloom, 1971; Carroll, 1963; Popham, 1969; Proger & Mann, 1973). With this type of evaluation, student performance is judged on some criterion of adequacy which is, in the student's perception, independent of his relative standing in a group of peers.

Two general types of criterion-referenced evaluation methods can be postulated. The typical and most often used type is that of a regulated criterion level. With this type, there is an established minimal level of competence which has been arrived at by some external authority. Such tests are constructed "to yield measurements that are directly interpretable in terms of specified performance standards," (Glasser & Nitko, 1971).

The second type of criterion-referenced evaluation can best be viewed as a negotiated form in which the acceptable level of competence is mutually established and agreed upon by both the external authority and the student being evaluated. This requires that the student become an active member of his own evaluation by assisting in the setting of a specific competence level for a particular learning task. It has been proposed that negotiated-criterion referenced

evaluation may enhance student motivation by actively involving the student in establishing the evaluation standards (Combs, 1976; Wilhelms, 1967).

Both norm-referenced and regulated criterion-referenced methods of evaluation can be viewed as involving external evaluation; that is, the student who is being evaluated has no input into the standards of evaluation. As long as the student's adequacy is being judged by external sources such as parents, teachers, and peers, then the student feels threatened and belittled (Wilhelms, 1967). The student begins to feel at the mercy of others -- someone to be manipulated (DeCharms, 1971).

The negotiated criterion-referenced system of evaluation involves both the student and the teacher in setting standards of acceptable performance, leading the student to perceive himself/herself as more in control of his/her evaluation. The student should thus feel better about performing toward those standards. Because the individual student is part evaluator, he/she could view "failure" as a temporary set-back to a goal rather than as being powerless to achieve the goal. It would appear that the failing individual would be more inclined to desire another chance to attempt success with the task in a negotiated situation.

Aspiration of Success

The concept of level of aspiration is defined as a means to study a subject's expectation or goal for some future attempt on a task (Diggory, 1966). Aspiration on tasks of criterion-referenced evaluation has generally indicated that aspiration rises as long as the individual's performance rises and lowers as performance lowers (Diggory & Morlock, 1964). The overall trend is for the level of aspiration to reflect the shape of the performance curve. Typically, the subject will set a fairly high initial level of aspiration, but will drastically reduce it after the first trial due to the feedback of falling quite short of his/her expectations. The level of aspiration then will continue to rise on each trial as long as performance increases. There should be significant differences between early adjacent trials as opposed to later trials with a negatively accelerated performance curve.

Typically, level of aspiration would differ under norm-referenced evaluation. A higher level of aspiration should be maintained than with criterion-referenced evaluation. This should occur because the individual under norm-references will soon realize that in order to progress in relative standing, he/she must perform at a higher rate than the norm,

and therefore, he/she should set a higher level of aspiration on the average than the criterion-referenced individuals.

Since the negotiated criterion-referenced group is more involved in the establishment of the criterion goal, it is expected that the aspiration of success for this group would be higher than for the regulated criterion-referenced group.

Academic Self-Concept

A number of studies indicate that self-concept and achievement in school is related (Alvord et al, 1967; Brookover et al, 1964; Centi, 1965; Diller, 1954; Gibby & Gibby, 1965; Morse, 1964; Marsh, et al., 1983).

The summary on self-evaluation research by Diggory (1966) indicates that, on criterion-referenced evaluations, an individual's self-evaluation will drop after failure.

Bloom (1976) presents research which indicates that, over a period of time, failures under a norm-referenced system lead to a progressive lowering of academic self-concept. Because of the difficulty in seeing objective improvement under norm-referenced conditions, it is believed that academic self-concept will be lower under norm-referencing than either regulated or negotiated criterion-

references.

Since the negotiated criterion-referenced group is more involved in the establishment of the criterion goal, it is expected that its academic self-concept would be higher than the academic self-concept of the regulated criterion-referenced group.

Test Anxiety

General anxiety is the anticipation of something unpleasant or painful whose source cannot be identified by the individual. In other words, the individual is apprehensive and has unpleasant feelings regarding future events, but cannot pinpoint the reason for them. Test anxiety, on the other hand, is anxiety about test taking. Test anxiety is found in some students who desire to achieve high. A low level of anxiety encourages effort to achieve; however, high test anxiety lowers achievement (Tobias, 1979). Furthermore, test anxiety increases as achieving high on a test becomes of greater importance (Deffenbacher, 1978).

High test-anxious students worry about the ill effects of not doing as well as they would like to do during a test. This self-oriented worry interferes with attending to and

concentrating on the test. Low test-anxious students do not worry; they concentrate only on the test (Wine, 1980). The lower problem-solving performance by a group of college students of high test anxiety was caused by an inability to recall information from memory and also by a failure to use an appropriate problem-solving strategy (Gross & Mastenbrook, 1980). The high anxiety interfered with recalling the needed information and with being able to use a known problem-solving strategy.

Since aspiration of success, academic self-concept, and test anxiety apparently have an effect on the evaluation of student performance, it becomes important to use an evaluation method that produces the least severe effects after failure. The question arises: "What effect does failure as evaluated by a norm-referenced measure, a regulated criterion-referenced measure, and a negotiated criterion-referenced measure have on aspiration of success, academic self-concept, and test anxiety?"

Most attention given to evaluation has been concentrated on the reliability and validity of a specified method in assessing ability. In the past, if the effects of an evaluation method on the student being evaluated were of importance, it was generally so in terms of "Does it

lead to improved cognitive or psychomotor performance?" Questions such as "How did the student feel about the process of learning?" have received less attention.

If Bloom (1976) is correct in his research, then approximately 20 to 25 percent of the variation in achievement can be accounted for in the student's affect which is one component of attitude, and that is too large a percentage to ignore.

The effects of the evaluation methods on attitudinal variables in this study could help to promote understanding of how evaluation effects students' interest and self-concept in relation to performance situations. Thus, this study is seen as continuing the research investigating the complex relationship between learner, task, performance, and attitude.

Method

Seventy-three college students from a large midwestern research university were randomly assigned to one of three testing groups: norm-referenced, regulated criterion-referenced, and negotiated criterion-referenced. Seventy-one subjects were female, two were male. Each group was given an analogies test over five repeated trials with three minutes

being the alotted time for each trial. This time limit was designed to enhance the likelihood of failure on all trials. Each analogies trial was a 20 item test derived from the Miller Analogies Test.

Before the first, third, and fifth trials, all subjects were asked to complete measures of test anxiety, academic self-concept, and aspiration of success that required an estimation by the student. The 0-10 scales that the students were asked to complete were preceded by: "Please circle the number that indicates the level of test anxiety you are experiencing at the present time." "Please circle the number that indicates your estimation of yourself as a student at the present time." And, "Please circle the number that indicates how well you expect to do on this series of analogies tests."

Subjects in the norm-referenced group were told that their scores would be based on a comparison with other students taking the test and that success was based on scoring at or above the 90th percentile. Subjects in the regulated criterion-referenced group were told that a score of 10 (out of 20) for each trial would be necessary in order to be considered proficient on the analogies tests.

The group selected as the negotiated criterion-

referenced group were given test packets that contained the following:

On the upcoming analogies tests you will be allowed to decide the score that you believe will indicate your proficiency.

The following scores represent average scores for students in each of three skill levels on a twenty item test:

7----------12----------16
(Low) (Average) (High)

Out of twenty items, what number do you think you should get correct in order to demonstrate proficiency?

_____(Enter a number between 7 and 20.)

Testing was completed within an hour. Subjects were then fully debriefed and told that none actually failed the tests and that the tests were in no way predictive of future academic performance.

Results

Means and standard deviations are displayed in Table 1. The data were analyzed using a split plot Analysis of Variance with Evaluation Method as the between subjects factor and time of administration (before Trials 1, 3 and 5) as the within subjects factor. Separate univariate analyses were conducted for academic self-concept, aspiration of

success, and test anxiety. Table 2 shows that for all dependent variables there was no significant main effect for Evaluation Method, nor was there a significant Method by Trials interaction. For all three analyses, the trials main effect was significant. Academic self-concept and aspiration of success declined from Pre-Trial to Trial 2 and then increased slightly. Test anxiety showed just the opposite pattern with an increase from Pre-Trial to Trial 2 followed by a slight decrease. (See Figure 1) Polynomial contrasts on the Trials factor showed the quadratic trend was significant for all three dependent measures.

Correlations among the dependent measures are displayed in Table 3. Prior to Trial 1 on the analogies test the correlation between academic self-concept and aspiration of success was .31 for the Negotiated Criterion-Referenced Group, .49 for the Regulated Criterion-Referenced Group ($p<.05$), but almost zero (-.06) for the Norm-Referenced Group. However, following Trial 2, the correlation was .67 ($p<.01$) for the Norm-Referenced Group and increased to .72 following Trial 4. The Regulated Criterion-Referenced Group consistently manifested the highest correlations between academic self-concept and test anxiety. Before Trial 1 correlation between test anxiety and aspiration of success was highest for the Norm-Referenced Group (-.43,

$p<.05$) and almost zero for the Negotiated Criterion-Referenced Group (-.02).

However, independent two-tailed t-tests for differences in correlations revealed no significant differences among groups.

Table 1
Means and Standard Deviations by Group

	N. C.	C. R.	N. R.
S. C.- Pre	7.91 (1.16)	8.00 (1.13)	7.08 (2.08)
S. C.- 2	4.00 (2.36)	4.63 (2.87)	4.23 (2.82)
S. C.- 4	5.22 (2.58)	5.29 (2.46)	4.69 (2.65)
ASP-Pre	6.78 (1.95)	7.05 (1.79)	6.58 (1.75)
ASP-2	3.70 (1.80)	3.33 (1.74)	3.39 (2.35)
ASP-4	4.44 (1.97)	4.25 (1.82)	4.15 (2.05)
TA-Pre	4.22 (2.84)	3.64 (2.94)	4.35 (3.56)
TA-2	5.91 (2.71)	5.13 (3.23)	5.00 (3.29)
TA-4	5.48 (2.66)	4.58 (2.90)	5.15 (3.33)
Trial 1	12.96 (3.53)	11.58 (3.65)	12.58 (3.84)
Trial 2	14.17 (4.39)	12.46 (3.74)	14.08 (2.90)
Trial 3	17.30 (2.65)	17.29 (3.50)	18.08 (2.77)
Trial 4	16.87 (3.21)	16.08 (4.34)	17.15 (3.16)
Trial 5	17.30 (3.27)	17.58 (2.90)	18.69 (2.00)

S. C. = Self-Concept
ASP = Aspiration for Success
TA = Test Anxiety

Table 2

ANOVA SUMMARY TABLE

Dependent Variable: Test Anxiety

Source	SS	df	MS	F
EM	23.843	2	11.921	.535
Error	1515.265	68	22.283	
Trial	60.289	2	30.144	9.385**
EM*Trial	8.666	4	2.167	.674
Error	436.846	136	3.212	

Dependent Variable: Academic Self-Concept

Source	SS	df	MS	F
EM	2.589	2	1.294	.170
Error	517.993	68	7.618	
Trial	432.182	136	1.830	118.065**
EM*Trial	3.346	4	.836	.457
Error	248.917	136	1.830	

Dependent Variable: Aspiration of Success

Source	SS	df	MS	F
EM	17.837	2	8.918	.737
Error	834.812	69	12.099	
Trial	441.182	2	220.640	103.822**
EM*Trial	7.465	4	1.866	.878
Error	293.275	138	2.125	

EM = Evaluation Method
** p <.001, * p <.05

Figure 1
Polynomial Contrasts of Trials

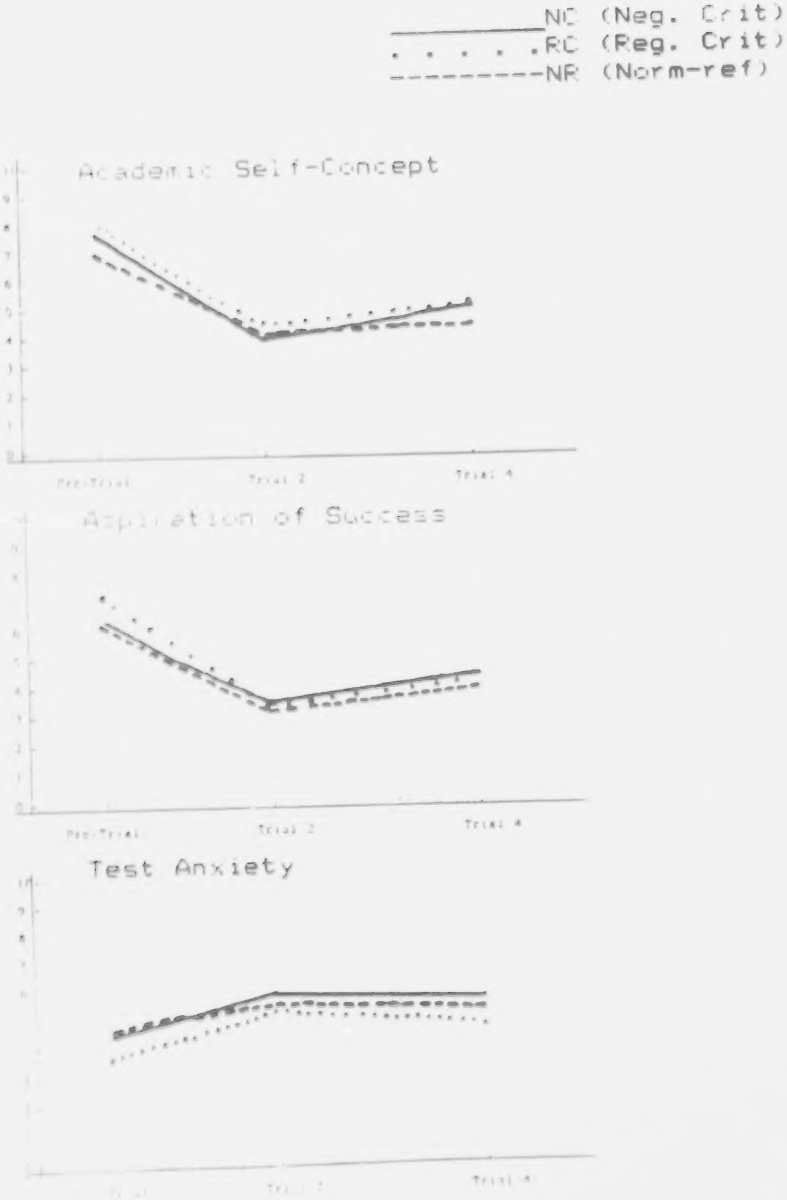

Table 3

CORRELATION MATRIX

	All Groups	NC	RC	NR
Pre-Trial	n=71	n=23	n=22	n=26
ASC * ASP	.18	.31	*.19	-.06
ASC * TA	-.11	-.08	-.30	-.02
ASP * TA	-.26*	-.02	-.28	-.43*
After Trial 2	n=73	n=23	n=24	n=26
ASC * ASP	.56**	.64**	.40+	.67**
ASC * TA	-.21	-.14	-.38+	-.08
ASP * TA	-.13	-.30	.05	-.20
After Trial 4	n=73	n=23	n=24	n=26
ASC * ASP	.65**	.54**	.68**	.72**
ASC * TA	-.26*	-.20	-.43*	-.12
ASP * TA	-.16	-.25	-.19	-.13

NC-Negotiated Criterion-Referenced
RC-Regulated Criterion-Referenced
NR-Norm-Referenced

ASC-Academic Self-Concept
ASP-Aspiration of Success
TA-Test Anxiety

* p<.05, ** p<.01, + p<.07

Discussion & Conclusions

Students in this study were given a limited amount of time to complete a series of difficult word analogies. The time limit was imposed so that students' academic self-concept, test anxiety, and aspiration of success could be studied under induced failure situations.

Comparisons involving treatment means showed that method of evaluation had no significant effect on students' academic self-concept, test anxiety or aspiration of success on the next trial. At pre-trial, students generally rated themselves as being successful students, expected to do well on the analogies test, and reported a moderate amount of test anxiety. However, following two series of 20 word analogy problems in which the stringent time limit guaranteed poor performance, not surprisingly, students reported their test anxiety level to increase and their academic self-concept and aspiration of success to decrease. Following two more grueling trials, students' self-reported test anxiety had moderated and their self-concept and aspiration of success showed a slight increase. The initial detrimental effect of failure on test anxiety, self-concept, and aspiration of success are consistent with the literature. However, the subsequent improvement on all three measures has a number of possible explanations. First, given that

students attempted more of the problems on the last three trials than on the first two, they may have "gotten the hang" of the analogies and felt that they were performing better. Alternatively, after four trials they may have realized the time limit was ridiculously short and just randomly filled in the answers. In this case the "noted improvement" may be due to the students' distancing themselves from any psychological impact their performance may have had. In other words, the tests were so difficult that students may have felt they were not really expected to perform well.

Although the method of evaluation apparently had no effect on the elevation of scores measuring academic self-concept, aspiration of success, and test anxiety, the relationships among these variables reflect some treatment effect. When students completed the pre-trial questionnaire, they had already been informed what level of performance was required in order for them to be considered proficient. For the norm-referenced group, this meant they had to do better than 90% of their peers; in other words, only 10% of them would be considered successful. Students in the regulated criterion-referenced group knew they needed 1/2 of the questions correct; whereas those in the negotiated criterion-referenced group had negotiated their own standard. The lack of relationship between academic self-concept and

aspiration of success for the norm-referenced group at pre-trial probably indicates that these students did not have any basis on which to judge their expected performance since it would be considered relative to others. Thus their self-concept would not provide them with a basis to judge. Students in both of the other groups had a number correct standard against which to gauge their expected performance; particularly the students in the regulated criterion-referenced group who knew what they needed to achieve and could use their own self-knowledge tc predict whether they would succeed or not.

After two failure experiences, however, the dramatic increase in correlation between achievement expectation and academic self-concept and aspiration of success for the norm-referenced group suggests that even without a number right criterion for success, these students based their success expectations on their past performances. In other words, although they did not know the criterion for success in terms of number right, their recent performance likely led them to believe they were not in the top ten percent.

The instruction, "Please circle the number that indicates how well you expect to do on this series of analogies tests," that preceded the aspiration of success scales was perhaps measuring expectation, not aspiration.

If the definition of aspiration is considered to be a part of a goal setting process and expectation is defined as predicting a likely outcome, this scale may have more aptly measured the students' expectations, rather than aspirations.

References

Alvord, D. J., & Glass, L. W. (1974). Relationships between academic achievement and self-concept. Science Education, 58, 175-179.

Astin, A. N., et al. (1967). National Norms for Entering College Freshmen 1966. Washington, D. C.: American Council of Education.

Block, J. H. (Ed.). (1971). Mastery Learning: Theory and Practice. New York: Holt, Rinehart and Winston.

Bloom, B. S. (1971). Learning for Mastery, Chapter 3 in B. S. Bloom, J. T. Hasting, & G. F. Madaus, Handbook on Formative and Sumative Evaluation of Student Learning. New York: McGraw-Hill Book Co.

Bloom, B. S. (1976). Human Characteristics and School Learning. New York: McGraw-Hill Book Co.

Brookover, W. B., et al. (1964). Self-concept of ability and school achievement. Sociology of Education, 37, 271-278.

Carroll, J. B. (1963). A Model of School Learning. Teachers College Record, 64, 723-733.

Centi, P. (1965). Self-perception of students and motivation. Catholic Education Review, 63, 307-319.

Combs, A. W. (1976). A contract method of evaluation, in Simon, S. B. and Bellance, J. A. Degrading the Grading Myths. Washington, D. C. : Association for Supervision and Curriculum Development.

De Charms, R. (1971). From pawns to origins: toward self-motivation, in Lesser, G. S. Psychology and Educational Practices. Glenview, Il : Scott, Foresman and Co.

Deffenbacher, J. L. (1978). Worry, emotionality, and task-generated interference in test anxiety: An empirical test of attentional theory. Journal of Educational Psychology, 70, 248-254.

Diggory, J. C., & Morlock, H. C., Jr. (1964). Level of aspiration, or probability of success? Journal of Abnormal and Social Psychology, 69, 282-289.

Diggory, J. C. (1966). Self-evaluation: Concepts and Studies. New York: John Wiley and Sons, Inc.

Diller, L. (1954). Conscious and unconscious self-attitudes after success and failure. Journal of Personality, 23, 1-12.

Gibby, R. G., & Gibby, R. C., Jr. (1965). The effects of stress resulting from academic failure. Journal of Clinical Psychology, 63, 307-319.

Glasser, R., & Nitko, A. J. (1971). Measurement in Learning and Instruction, in Thorndike, R. L., (Ed.), Educational Measurement. Washington, D. C.: American Council on Education.

Gross, T. E., & Mastenbrook, M. (1980). Examination of the effects of state anxiety on problem-solving efficiency under high and low memory conditions. Journal of Educational Psychology, 72, 605-609.

Marsh, H. W., Smith, I.d., Barnes, J., & Butler, S. (1980) Self-concept: Reliability, stability, dimensionality, validity, and the measurement of change. Journal of Educational Psychology, 75, 772-790.

Morse, L. F. (1964). Self-concept in the school setting. Childhood Education, 41, 195-198.

Peckham, P. D., & Roe, M. D. (1977). The effects of frequent testing. Journal of Research and Development in Education, 10, 40-50.

Popham, W. J., & Husek, T. R. (1969). Implications of criterion referenced measurement. Journal of Educational Measurements, 6, 1-9.

Popham, W. J. (1971). Criterion-referenced Measurement: An Introduction. Englewood Cliffs, N. J.: Educational Technology Publications.

Proger, B. B., & Mann, L. (1973). Criterion-referenced measurement: the world gray versus black and white. Journal of Learning Disabilities, 6, 18-30.

Tobias, S. (1979). Anxiety research in educational psychology. Journal of Educational Psychology, 71, 573-

Wilhelms, F. T. (1967). Evaluation as Feedback and Guide. Washington, D. C.: Association for Supervision and Curriculum Development.

Wine, J. D. (1980). Cognitive-attentional theory of test anxiety. In I. G. Sarason (Ed.), Test anxiety: Theory, research and application. Hillside, N. J.:P Erlbaum.